Les 3 Verwachtingswaarde en spreiding

3.1 Stochasten

In een paar voorbeelden hebben we al gezien dat we bij een experiment vaak niet zo zeer in een enkele uitkomst geïneteresseerd zijn, maar bijvoorbeeld wel in het aantal uitkomsten van een zeker soort. Zo willen we bij een steekproef weten, hoeveel stukken defect zijn, maar niet of nu het eerste of laatste stuk defect is.

Vaak zijn de uitkomsten waarin we geïneteresseerd zijn veel eenvoudiger dan de uitkomstenruimte zelf, bijvoorbeeld kijken we naar het aantal k van defecte stukken in plaats van alle combinaties van m testresultaten, waarvan k negatief zijn. We kunnen dus zeggen, dat we verschillende uitkomsten die een zekere eigenschap gemeenschappelijk hebben in een cluster samenvatten, Zo'n eigenschap laat zich door een functie

$$X : \Omega \to \mathbb{R}, \quad \omega \mapsto X(\omega)$$

beschrijven, die aan elk element ω van de uitkomstenruimte een waarde $X(\omega)$ toekent. Zo'n functie X noemen we een *random variable* (in het Engels), een *stochastische variabele*, een *kansvariabele* of kort een *stochast*.

In het voorbeeld van de kwaliteitsproef is de stochast dus de functie die aan een rij van testresultaten het aantal negatieve (of positieve) resultaten toekent.

Een ander voorbeeld is het dobbelen met twee dobbelstenen: Als we alleen maar in de som van de geworpen getallen geïneteresseerd zijn, nemen we als stochast de functie $X(\omega_1, \omega_2) := \omega_1 + \omega_2$.

Het belangrijke aan de stochasten is, dat we makkelijk een kansverdeling hiervoor kunnen definiëren: De kans $P(X = x)$ dat de stochast de waarde x aanneemt definiëren we door

$$P(X = x) := \sum_{X(\omega)=x} P(\omega)$$

dus we tellen gewoon de kansen voor alle elementen van Ω op, waar de stochast de waarde x oplevert.

Onbewust hebben we al eerder stochasten op deze manier gebruikt, bijvoorbeeld voor het uitrekenen van de kans dat we met twee dobbelstenen een som van 5 werpen.

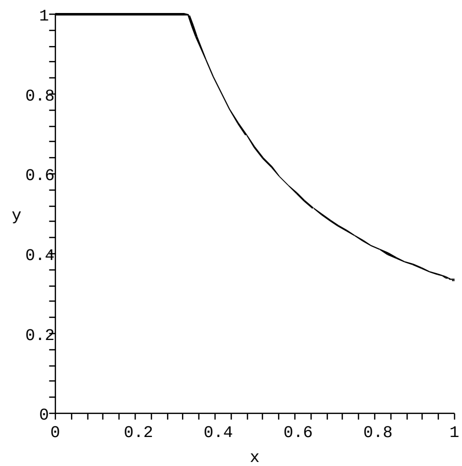
Voor continue kansverdelingen gaat de som over de uitkomsten met $X(\omega) = x$ over in een integraal. Omdat de kans op een enkele uitkomst steeds 0 is, wordt hier de kans bepaald, dat de stochast X een waarde onder een gegeven grens aanneemt. Voor een continue kansverdeling met dichtheidsfunctie $f(x)$ krijgen we:

$$P(X \leq x) = \int_{X(t)\leq x} f(t)\, dt.$$

Meestal zijn continue stochasten door hun eigen dichtheidsfunctie aangegeven, er geldt dan

$$P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt.$$

Voorbeeld: Stel we hebben een randomgenerator die toevalsgetallen tussen 0 en 1 volgens de uniforme verdeling voortbrengt. We vragen ons af, wat de kans is dat het product van twee opeenvolgende van die toevalsgetallen kleiner is dan een grens $0 \leq a \leq 1$. De stochast die bij dit probleem hoort is $X(x, y) := x \cdot y$ en omdat we het met de uniforme verdeling te maken hebben, moeten we alleen maar de oppervlakte van het gebied $G = \{(x, y) \in \mathbb{R}^2 \mid x \cdot y \leq a\}$ bepalen. Als $x \leq a$ kan y elke waarde tussen 0 en 1 hebben, maar voor $x \geq a$ hebben we $y \leq \frac{a}{x}$ nodig. De volgende schets laat dit (voor $a = \frac{1}{3}$) zien:



Met behulp van een eenvoudige integratie kunnen we de kansverdeling van deze stochast ook expliciet bepalen, er geldt:

$$P(X \leq a) = \int_0^a dx + \int_a^1 \frac{a}{x}\, dx = a + a(\log(1) - log(a)) = a(1 - \log(a)).$$

Voor $a = 0.5$ is deze kans bijvoorbeeld $P(X \leq 0.5) \approx 0.85$ en pas voor $a < 0.187$ is $P(X \leq a) < 0.5$.

3.2 Verwachtingswaarde

Als we in het casino roulette gaan spelen, zijn we er niet in geïnteresseerd of we in het eerste of laatste spel winnen of verliezen en ook niet hoe vaak we winnen of verliezen. Eigenlijk willen we alleen maar weten of we kunnen verwachten dat we aan het eind van de dag met een winst naar huis komen. Als we N keer spelen en bij elke keer 10€ op rood zetten, dan is bij elk spel de kans dat we 10€ winnen gelijk aan $\frac{18}{37}$, want er zijn 18 rode en 18 zwarte getallen en de groene 0. De kans dat we de 10€ verliezen is dus $\frac{19}{37}$. Als we heel vaak spelen, kunnen we verwachten dat we $\frac{18 \cdot N}{37}$ keer winnen en $\frac{19 \cdot N}{37}$ keer verliezen. Dit betekend dat we een verlies van $N \cdot \frac{1}{37} \cdot 10€$ kunnen verwachten.

Uit het perspectief van het casino is dit natuurlijk heel wenselijk. Omdat alle winsten alleen maar op de getallen 1 t/m 36 zijn gebaseerd (dus als je op de 3 getallen $4, 5, 6$ zet maak je een winst van 12 keer je inzet) heeft de groene 0 het effect dat het casino gemiddeld een zevenendertigste van alle inzetten wint.

In het voorbeeld van het roulette spel hebben we een stochast gebruikt die het bedrag van de winst of verlies aangeeft. Waar we in geïnteresseerd zijn is de gemiddelde winst die we per spel zullen maken. Dit is het gemiddelde van de mogelijke waarden van de stochast, waarbij elke waarde met zijn kans gewogen wordt. Wat we zo krijgen is de winst die we per spel gemiddeld verwachten, en daarom noemen we dit ook de *verwachtingswaarde*.

Algemeen definiëren we voor een stochast X de verwachtingswaarde $E(X)$ (de E staat voor het Engelse *expectation*) door

$$E(X) := \sum_{x \in X} x \cdot P(X = x) = \sum_{x \in X} x \cdot \Big(\sum_{X(\omega)=x} P(\omega)\Big) = \sum_{\omega \in \Omega} X(\omega) P(\omega).$$

Voor een stochast X met continue kansverdeling is de verwachtingswaarde met behulp van zijn dichtheidsfunctie $f(x)$ analoog gedefinieerd door de integraal

$$E(X) := \int_{-\infty}^{\infty} x \cdot f(x)\, dx.$$

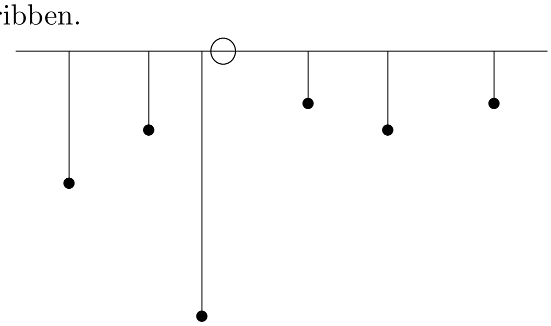
Merk op dat we van een continu verdeelde stochast door samenvatten van de waarden in een deelinterval naar een discreet verdeelde stochast kunnen komen:

Er geldt $P(X \in [x, x+\delta]) = \int_x^{x+\delta} f(t)\, dt$ en voor kleine δ kunnen we aannemen dat $f(t)$ op het interval $[x, x+\delta]$ bijna constant is, dit geeft

$$P(X \in [x, x+\delta]) \approx \delta \cdot f(x).$$

Als we nu de reële lijn in stukjes $[i \cdot \delta, (i+1) \cdot \delta]$ van lengte δ onderverdelen en de uitkomsten $x \in [i \cdot \delta, (i+1) \cdot \delta]$ tot de uitkomst $x = i \cdot \delta$ samenvatten, hebben we alleen maar nog de discrete verzameling $\{i \cdot \delta \mid i \in \mathbb{Z}\}$ van uitkomsten. Voor deze *gediscretiseerde* stochast is de verwachtingswaarde $\sum_{i \in \mathbb{Z}, x = i \cdot \delta} x \cdot P(X \in [x, x+\delta]) \approx \sum_{i \in \mathbb{Z}, x = i \cdot \delta} x \cdot \delta \cdot f(x)$ en dit is juist de discrete benadering van de integraal $\int_{-\infty}^{\infty} x \cdot f(x)\, dx = E(X)$.

We kunnen de verwachtingswaarde aanschouwelijk zien als het evenwichtspunt van een balk (oneindig lang, zonder gewicht), waar we in het punt x een gewicht van massa $P(x)$ aan hangen. Het evenwichtspunt is dan juist het punt $E(X)$. In het volgende plaatje zijn de gewichten gerepresenteerd door de lengten van de verticale ribben.



Een aantal belangrijke elementaire eigenschappen van de verwachtingswaarde kunnen we meteen uit de definitie aflezen. Als X en Y stochasten zijn, dan geldt:

(i) $E(X+Y) = E(X) + E(Y)$, dus de som van de verwachtingswaarden van twee stochasten is de verwachtingswaarde van de som van de stochasten.

(ii) $E(\alpha X) = \alpha E(X)$.

(iii) $X(\omega) \geq Y(\omega)$ voor alle $\omega \in \Omega \Rightarrow E(X) \geq E(Y)$.

Als we in (i) voor Y de constante stochast $Y(\omega) = c$ nemen, volgt hieruit dat een verschuiving van de stochast om c ook de verwachtingswaarde om c verschuift (omdat de constante stochast verwachtingswaarde c heeft). We kunnen dus een stochast door aftrekken van zijn verwachtingswaarde altijd zo verschuiven dat hij verwachtingswaarde 0 heeft:

$$X_0 := X - E(X) \Rightarrow E(X_0) = E(X - E(X)) = E(X) - E(X) = 0.$$

We gaan nu de verwachtingswaarden van de belangrijkste kansverdelingen berekenen.

Binomiale verdeling

We hebben $P(X = k) = b(m, p; k) = \binom{m}{k} p^k (1-p)^{m-k}$, dus:

$$\begin{aligned}
E(X) &= \sum_{k=0}^{m} k \binom{m}{k} p^k (1-p)^{m-k} = \sum_{k=0}^{m} k \, \frac{m!}{k!(m-k)!} \, p^k (1-p)^{m-k} \\
&= m \cdot p \cdot \sum_{k=1}^{m} \frac{(m-1)!}{(k-1)!(m-k)!} \, p^{k-1} (1-p)^{m-k} \\
&= m \cdot p \cdot \sum_{k=0}^{m-1} \binom{m-1}{k} p^k (1-p)^{m-1-k} \\
&= m \cdot p \cdot \sum_{k=0}^{m-1} b(m-1, p; k) = m \cdot p.
\end{aligned}$$

In de laatste stap hebben we hierbij gebruik van het feit gemaakt, dat de som over de kansen $b(m-1, p; k)$ voor alle waarden van k de totale kans 1 oplevert. De verwachtingswaarde van de binomiale verdeling is dus $m \cdot p$ en dit is precies het verwachte aantal van gunstige uitkomsten als we m pogingen doen bij een kans van p voor een gunstige uitkomst.

Hypergeometrische verdeling

We hebben $P(X = k) = h(n, m, s; k) = \frac{\binom{s}{k} \cdot \binom{n-s}{m-k}}{\binom{n}{m}}$, en er geldt: $k \cdot \binom{s}{k} = k \cdot \frac{s!}{k!(s-k)!} = s \cdot \frac{(s-1)!}{(k-1)!(s-k)!} = s \cdot \binom{s-1}{k-1}$ en $\binom{n}{m} = \frac{n!}{m!(n-m)!} = \frac{n}{m} \cdot \frac{(n-1)!}{(m-1)!(n-m)!} =$

$\frac{n}{m} \cdot \binom{n-1}{m-1}$. Hieruit volgt:

$$E(X) = \sum_{k=0}^{m} k \frac{\binom{s}{k} \cdot \binom{n-s}{m-k}}{\binom{n}{m}} = \sum_{k=1}^{m} \frac{s\binom{s-1}{k-1} \cdot \binom{n-s}{m-k}}{\frac{n}{m}\binom{n-1}{m-1}} = m\frac{s}{n} \sum_{k=1}^{m} \frac{\binom{s-1}{k-1} \cdot \binom{n-s}{m-k}}{\binom{n-1}{m-1}}$$
$$= m\frac{s}{n} \sum_{k=0}^{m-1} \frac{\binom{s-1}{k} \cdot \binom{n-s}{m-1-k}}{\binom{n-1}{m-1}} = m\frac{s}{n} \sum_{k=0}^{m-1} h(n-1, m-1, s; k) = m\frac{s}{n}.$$

In de stap naar de laatste regel hebben hierbij k door $k+1$ verplaatst, de som die voor k van 1 tot m loopt, loopt voor $k+1$ van 0 tot $m-1$. In de laatste stap loopt de som over de kansen $h(n-1, m-1, s; k)$ voor alle waarden van k, dus is deze som gelijk aan 1. Het resultaat hadden we ook intuïtief kunnen afleiden, want de kans om bij een greep een van de s slechte stukken uit de totale n stukken te pakken is $\frac{s}{n}$, en als we m keer grijpen zouden we gemiddeld $m\frac{s}{n}$ slechte stukken verwachten.

Poisson-verdeling

We hebben $P(X = k) = po_\lambda(k) = \frac{\lambda^k}{k!}e^{-\lambda}$ en maken gebruik van de gelijkheid $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^\lambda$:

$$E(X) = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda \cdot e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda \cdot e^{-\lambda} \cdot e^\lambda = \lambda.$$

Ook hier vinden we het verwachte resultaat, omdat de Poisson-verdeling de limiet van de binomiale verdeling is als $p \to 0$ gaat en $m \cdot p = \lambda$ constant is.

Uniforme verdeling

We hebben $P(X = x) = \frac{1}{b-a}$ als $a \leq x \leq b$ en 0 anders, dus

$$E(X) = \int_a^b x\frac{1}{b-a}\, dx = \frac{1}{2(b-a)}(b^2 - a^2) = \frac{1}{2}(a+b).$$

De verwachtingswaarde is dus het middelpunt van het interval waarop de dichtheidsfunctie niet 0 is.

Exponentiële verdeling

We nemen aan dat we de dichtheidsfunctie zo hebben verschoven dat de beginwaarde $c = 0$ is. Dan is $f(x) = \lambda e^{-\lambda x}$ als $x \geq 0$ en $f(x) = 0$ anders. Dit geeft

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x}\, dx = -x\lambda e^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x}\, dx = -\frac{1}{\lambda}e^{-\lambda x}\Big|_0^\infty = \frac{1}{\lambda}$$

(merk op dat we hierbij gebruiken dat $\lim_{x\to\infty} xe^{-x} = 0$ is). Ook hier is het resultaat voor de verwachtingswaarde plausibel, want als λ groter wordt, gaat de functie $f(x)$ sneller naar nul en moeten we dus een kleinere verwachtingswaarde krijgen.

Normaalverdeling

In dit geval kunnen we de verwachtingswaarde zonder enig rekenwerk bepalen. Als we de dichtheidsfunctie $f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ zo verschuiven dat $\mu = 0$ is, is de functie symmetrisch ten opzichte van de y-as en dan is $E(X) = 0$. De verwachtingswaarde voor de algemene normaalverdeling is dus μ en dit is ook geen verrassing omdat de dichtheidsfunctie juist zo gemaakt is.

3.3 Spreiding

Als we de verwachtingswaarde van een stochast kennen, weten we wat we op lange termijn gemiddeld kunnen verwachten. Maar vaak willen we toch iets meer weten, bijvoorbeeld hoe ver de daadwerkelijke uitkomsten van de verwachtingswaarde verwijderd zijn. Als we namelijk een stochast X zo verschuiven dat de verwachtingswaarde 0 is, dan heeft ook de stochast αX verwachtingswaarde 0, maar voor $\alpha > 1$ zijn de enkele uitkomsten verder van de verwachtingswaarde verwijderd.

In het model van de balk met gewichten kunnen we het verschil tussen de stochasten X en αX duidelijk zien. Als de gewichten dicht bij het evenwichtspunt zijn, kunnen we de balk makkelijk om dit punt draaien. Als we nu bijvoorbeeld naar de stochast $10 \cdot X$ kijken, worden de afstanden van het evenwichtspunt met 10 vermenigvuldigd. Nu hebben we meer kracht nodig om de balk te draaien. Dit ligt eraan dat het traagheidsmoment van de balk groter geworden is, dit is namelijk gegeven als als de som over $m \cdot r^2$ waarbij m de massa in een punt is die afstand r van het draaipunt heeft. Als we het traagheidsmoment naar de stochast vertalen wordt dit

$$Var(X) := \sum_{x \in X} (x - E(X))^2 \cdot P(X = x) = E((X - E(X))^2)$$

en dit noemen we de *variantie* of *spreiding* van X. De variantie is dus de verwachtingswaarde van de kwadratische afstand van de stochast van zijn verwachtingswaarde en is dus een maat ervoor hoe dicht de waarden van een stochast bij de verwachtingswaarde liggen.

Vaak wordt in plaats van de variantie de wortel uit de variantie als maat voor de afwijkingen gebruikt, omdat deze lineair met de stochast verandert (d.w.z. als we X met een factor α vermenigvuldigen, wordt ook de wortel uit de variantie met α vermenigvuldigt). We definiëren dus

$$\sigma_X := \sqrt{Var(X)}$$

en noemen dit de *standaardafwijking* van X.

Voorbeeld: Bij het werpen van een dobbelsteen is de verwachtingswaarde $E(X) = \sum_{k=1}^{6} k \cdot \frac{1}{6} = \frac{7}{2}$. De variantie is dan $Var(X) = \sum_{k=1}^{6}(k - \frac{7}{2})^2 \cdot \frac{1}{6} = \frac{35}{12}$ en de standaardafwijking $\sigma_X = \sqrt{\frac{35}{12}} \approx 1.7$.

We hebben boven opgemerkt dat de variantie van een stochast aangeeft hoe sterk de uitkomsten van de verwachtingswaarde afwijken. Deze samenhang tussen verwachtingswaarde en spreiding kunnen we heel expliciet aangeven, namelijk in de *Ongelijkheid van Chebyshev.* Hierbij maken we een afschatting voor de kans dat een uitkomst een grotere afstand dan $a > 0$ van de verwachtingswaarde $E(X)$ heeft.

Volgens de definitie berekenen we de variantie door $Var(X) = \sum_{x \in X}(x - E(X))^2 \cdot P(X = x)$. Als we de som beperken tot de waarden van x met $|x - E(X)| \geq a$, krijgen we

$$Var(X) \geq \sum_{|x-E(X)| \geq a} (x - E(X))^2 \cdot P(X = x) \geq \sum_{|x-E(X)| \geq a} a^2 \cdot P(X = x)$$

en dit is juist $a^2 \cdot P(|X - E(X)| \geq a)$. We hebben dus bewezen:

Ongelijkheid van Chebyshev: Voor een stochast X met verwachtingswaarde $E(X)$ en variantie $Var(X)$ geldt voor elke $a > 0$ de ongelijkheid

$$P(|X - E(X)| \geq a) \leq \frac{1}{a^2} Var(X).$$

Als voorbeeld kunnen we met de ongelijkheid van Chebyshev eens afschatten, wat de kans op het dobbelen van een zes is. We hebben boven gezien dat de verwachtingswaarde bij het dobbelen $\frac{7}{2}$ en de variantie $\frac{35}{12}$ is. De afstand tussen een 6 en de verwachtingswaarde $\frac{7}{2}$ is $\frac{5}{2}$ en volgens de ongelijkheid van Chebyshev geldt $P(|X - E(X)| \geq \frac{5}{2}) \leq \frac{4}{25} \cdot \frac{35}{12} = \frac{7}{15} \approx 0.467$. Omdat deze kans ook het dobbelen van een 1 insluit, moeten we nog door twee delen en schatten de kans op een 6 dus met 23.3% (naar boven) af. Natuurlijk weten we dat de kans in feite $\frac{1}{6} = 16.7\%$ is en dit laat zien dat de afschatting niet eens zo slecht is.

In de statistiek wordt vaak als vuistregel de zo genoemde 2σ-regel gebruikt: Voor een stochast X met standaardafwijking σ_X liggen meestal 95% van de gebeurtenissen in het interval $(E(X) - 2\sigma_X, E(X) + 2\sigma_X)$. De ongelijkheid van Chebyshev geeft aan dat dit interval minstens 75% van de gebeurtenissen bevat, maar voor de meeste kansverdelingen (in het bijzonder voor de normaalverdeling) geldt de sterkere uitspraak van de 2σ-regel.

Naast de ongelijkheid van Chebyshev kunnen we een aantal verdere belangrijke eigenschappen voor de variantie van een stochast X meteen uit de definities afleiden:

(i) $Var(X) = 0$ dan en slechts dan als $X = c$ constant is.

(ii) $Var(\alpha X) = \alpha^2 Var(X)$ en $\sigma_{\alpha X} = \alpha \cdot \sigma_X$.

(iii) $Var(X+c) = Var(X)$, dus zo als we dit zouden verwachten is de variantie onafhankelijk van een verschuiving van de stochast.

(iv) $Var(X) = E(X^2) - E(X)^2$, want:

$$\begin{aligned} Var(X) &= \sum_{x \in X} (x - E(X))^2 \cdot P(X = x) \\ &= (\sum_{x \in X} x^2 \cdot P(X = x)) - 2E(X)(\sum_{x \in X} x \cdot P(X = x)) + E(X)^2 \\ &= E(X^2) - 2E(X) \cdot E(X) + E(X)^2 = E(X^2) - E(X)^2. \end{aligned}$$

Dit is in veel gevallen een handige formule om de variantie van een stochast uit te rekenen.

Vaak is het nuttig een stochast zo te normeren dat hij verwachtingswaarde 0 en variantie 1 heeft. Dit kunnen we met behulp van (ii) en (iii) makkelijk bereiken, want voor $X_0 := \frac{X-E(X)}{\sigma_X}$ geldt $E(X_0) = \frac{1}{\sigma_X}(E(X) - E(X)) = 0$ en $Var(X_0) = Var(\frac{X}{\sigma_X}) = \frac{1}{\sigma_X^2}Var(X) = 1$.

We gaan nu ook de varianties van de meest belangrijke kansverdelingen berekenen.

Binomiale verdeling

Dit pakken we met de formule $Var(X) = E(X^2) - E(X)^2$ aan:

$$\begin{aligned} E(X^2) &= \sum_{k=0}^{m} k^2 \binom{m}{k} p^k (1-p)^{m-k} \\ &= m \cdot p \cdot \sum_{k=1}^{m} k \frac{(m-1)!}{(k-1)!(m-k)!} p^{k-1} (1-p)^{m-k} \\ &= m \cdot p \cdot \sum_{k=0}^{m-1} (k+1) \binom{m-1}{k} p^k (1-p)^{m-1-k}. \end{aligned}$$

De som $\sum_{k=0}^{m-1}(k+1)\binom{m-1}{k}p^k(1-p)^{m-1-k}$ is de verwachtingswaarde van de verschoven stochast $X+1$ voor de parameter $m-1$, dus is de waarde hiervan $(m-1)p+1$. We hebben dus $E(X^2) = mp((m-1)p+1) = mp(mp+(1-p))$ en dus

$$Var(X) = E(X^2) - E(X)^2 = mp(mp + (1-p)) - (mp)^2 = mp(1-p).$$

Hypergeometrische verdeling

Dit is een beetje omslachtig om uit te werken, dus geven voor de volledigheid alleen maar het resultaat aan. Voor een stochast X met $P(X=k) = h(n,m,s;k)$ geldt

$$Var(X) = m\frac{s}{n}(1 - \frac{s}{n})\frac{n-m}{n-1}.$$

Als n veel groter is dan m geldt $\frac{n-m}{n-1} \approx 1$ en met $p = \frac{s}{n}$ wordt de variantie van de hypergeometrische verdeling dan benadert door de variantie van de binomiale verdeling.

Poisson-verdeling

We gebruiken weer de formule $\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{\lambda}$. Er geldt:

$$
\begin{aligned}
E(X^2) &= \sum_{k=0}^{\infty} k^2 \,\frac{\lambda^k}{k!}\, e^{-\lambda} = \sum_{k=1}^{\infty} k \,\frac{\lambda^k}{(k-1)!}\, e^{-\lambda} = \sum_{k=1}^{\infty}((k-1)+1)\,\frac{\lambda^k}{(k-1)!}\, e^{-\lambda} \\
&= (\sum_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!}\, e^{-\lambda}) + (\sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!}\, e^{-\lambda}) \\
&= \lambda^2 e^{-\lambda}(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}) + \lambda e^{-\lambda}(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}) = \lambda^2 + \lambda.
\end{aligned}
$$

We hebben dus

$$
Var(X) = E(X^2) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.
$$

Dit hadden we ook uit de variantie voor de binomiale verdeling kunnen gokken, want de Poisson-verdeling is de limiet voor $p \to 0$ met $mp = \lambda$ en bij deze limiet gaat $mp(1-p)$ naar $mp = \lambda$.

Uniforme verdeling

Er geldt

$$
E(X^2) = \int_a^b x^2 \frac{1}{b-a}\; dx = \frac{1}{3(b-a)}(b^3 - a^3) = \frac{1}{3}(a^2 + ab + b^2)
$$

dus hebben we

$$
Var(X) = E(X^2) - E(X)^2 = \frac{1}{3}(a^2 + ab + b^2) - \frac{1}{4}(a^2 + 2ab + b^2) = \frac{1}{12}(a-b)^2.
$$

Exponentiële verdeling

Er geldt

$$
E(X^2) = \int_0^{\infty} x^2 \lambda e^{-\lambda x}\; dx = -x^2 \lambda e^{-\lambda x}\Big|_0^{\infty} + 2\int_0^{\infty} x e^{-\lambda x}\; dx = 2\int_0^{\infty} x e^{-\lambda x}\; dx
$$

en dit is gelijk aan $\frac{2}{\lambda}E(X) = \frac{2}{\lambda^2}$. We hebben dus

$$
Var(X) = E(X^2) - E(X)^2 = \frac{2}{\lambda^2} - (\frac{1}{\lambda})^2 = \frac{1}{\lambda^2}.
$$

Normaalverdeling

Dit is iets ingewikkelder te berekenen maar de parameters in de normaalverdeling zijn zo gekozen dat σ^2 de variantie aangeeft en dus σ de standaardafwijking.

3.4 Covariantie en correlatie

Het is iets moeilijker om iets over de variantie van de som van twee stochasten te zeggen dan dit bij de verwachtingswaarde het geval was. We hebben

$$\begin{aligned}
Var(X+Y) &= E((X+Y)^2) - (E(X+Y))^2 \\
&= E(X^2 + 2X \cdot Y + Y^2) - (E(X) + E(Y))^2 \\
&= E(X^2) + 2E(X \cdot Y) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2 \\
&= E(X^2) - E(X)^2 + E(Y^2) - E(Y)^2 + 2E(X \cdot Y) - 2E(X)E(Y) \\
&= Var(X) + Var(Y) + 2(E(X \cdot Y) - E(X) \cdot E(Y)).
\end{aligned}$$

We noemen $E(X \cdot Y) - E(X) \cdot E(Y)$ de *covariantie* van X en Y en noteren dit met $Cov(X,Y)$. Er geldt dus

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

en dit betekent dat de covariantie aangeeft hoe sterk de variantie van de som van twee stochasten afwijkt van de som van de varianties.

De covariantie laat zich ook beschrijven als de verwachtingswaarde van het product van $(X - E(X))$ en $(Y - E(Y)$, want:

$$\begin{aligned}
E((X - E(X)) \cdot (Y - E(Y))) &= E(X \cdot Y - E(X)Y - E(Y)X - E(X)E(Y)) \\
&= E(X \cdot Y) - E(E(X)Y) - E(E(Y)X) + E(E(X)E(Y)) \\
&= E(X \cdot Y) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\
&= E(X \cdot Y) - E(X)E(Y) = Cov(X,Y),
\end{aligned}$$

dus hebben we

$$Cov(X,Y) = E((X - E(X))(Y - E(Y)).$$

We zullen in de volgende les uitgebreid bediscussiëren wat het betekent dat twee stochasten *onafhankelijk* zijn, maar intuïtief zou men al zeggen, dat de uitkomst van de ene stochast de uitkomst van de andere niet mag beïnvloeden. We zullen twee stochasten X en Y onafhankelijk noemen, als de kans $P(X = x, Y = y)$ op de gecombineerde uitkomst $X = x$ en $Y = y$ gelijk is aan het product $P(X = x) \cdot P(Y = y)$ van de kansen op de aparte uitkomsten en als dit voor alle paren (x, y) geldt.

Stel nu dat X en Y onafhankelijke stochasten zijn, dan geldt:

$$\begin{aligned}
E(X \cdot Y) &= \sum_{(x,y) \in X \times Y} x \cdot y \cdot P(X = x, Y = y) \\
&= \sum_{(x,y) \in X \times Y} x \cdot y \cdot P(X = x) \cdot P(Y = y) \\
&= (\sum_{x \in X} x \cdot P(X = x))(\sum_{y \in Y} y \cdot P(Y = y)) = E(X) \cdot E(Y).
\end{aligned}$$

We hebben dus gezien:

Voor onafhankelijke stochasten X en Y geldt $E(X \cdot Y) = E(X) \cdot E(Y)$, dus $Cov(X, Y) = 0$ en dus $Var(X + Y) = Var(X) + Var(Y)$.

Waarschuwing: De omkering hiervan geldt niet. Twee stochasten kunnen covariantie 0 hebben zonder onafhankelijk te zijn.

We hebben gezien dat de covariantie $Cov(X, Y)$ in zekere zin en maat voor de afhankelijkheid van X en Y is. Er laat zich aantonen dat $|Cov(X, Y)| \leq \sigma_X \sigma_Y$ is, dus de covariantie van twee stochasten is begrensd door het product van de standaardafwijkingen van de stochasten. Met behulp van de standaardafwijkingen kunnen we dus de covariantie op waarden tussen -1 en 1 normeren. We noemen

$$\rho_{X,Y} := \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

de *correlatiecoëfficiënt* van X en Y. De waarde van de correlatiecoëfficiënt ligt tussen -1 en 1 de waarde $\rho_{X,Y} = -1$ treedt alleen maar op voor $Y = -\alpha X + \beta$ met $\alpha > 0$, de waarde $\rho_{X,Y} = 1$ alleen maar voor $Y = \alpha X + \beta$ met $\alpha > 0$. Precies gezegd geeft de correlatiecoëfficient dus aan, in hoeverre de stochasten X en Y *lineair* van elkaar afhangen, d.w.z. hoe goed zich Y door $\alpha X + \beta$ laat benaderen. Voor $\rho_{X,Y} > 0$ spreekt men van *positieve afhankelijkheid* voor $\rho_{X,Y} < 0$ van *negatieve afhankelijkheid*.

Belangrijke begrippen in deze les

- stochasten
- verwachtingswaarde
- variantie, standaardafwijking
- covariantie, correlatiecoëfficiënt

Opgaven

11. Bij een spel met een dobbelsteen win je n€ als je n dobbelt en n even is en je verliest n€ als n oneven is. Wat is de verwachtingswaarde van je winst/verlies.
12. Bij het skaat spel krijg je 10 kaarten uit een kaartspel met 32 kaarten (8 soorten, 4 kleuren). Wat is de verwachtingswaarde voor het aantal boeren dat je krijgt?
13. In een loterij heb je 70% nieten en 30% winnende lotjes. Iemand beslist zo lang lotjes te kopen tot dat hij een winnende lot krijgt, maar hooguit vijf keer. Wat kan hij voor een uitgave verwachten, als een lot 2€ kost?
14. Je koopt een nieuwe speelautomaat voor je kroeg. In de automaat draaien twee onafhankelijke wielen die in tien even grote segmenten zijn opgedeeld en volgens een gelijkverdeling in een van de segmenten stoppen. De segmenten hebben de nummers 1 t/m 10. Een speler heeft alleen maar de volgende winstmogelijkheden (bij alle andere uitkomsten verliest hij zijn inzet):
    - Als beide wielen 10 tonen wint hij 5€.

- Als beide wielen hetzelfde getal maar niet 10 tonen wint hij 2€.
- Als precies een van de wielen 10 toont wint hij 1€.

Je wilt natuurlijk winst met je automaat maken. Wat is de minimale inzet die je per spel moet vragen om een winst te kunnen verwachten?

15. Twee tennissters A en B spelen vaker tegen elkaar en gemiddeld wint A 60% van de sets. De speelsters ontmoeten elkaar op een toernooi in een best-of-five match (dus wie het eerst drie sets wint heeft gewonnen).

    (i) Wat zijn de kansen dat A in 3, 4 of 5 sets wint? Hoe zit het met B? Wat is de kans dat B überhaupt wint?

    (ii) Bereken de verwachtingswaarde voor het aantal sets die het match duurt.

    (iii) Bereken apart de verwachtingswaarden voor het aantal sets in het geval dat A wint en dat B wint.

    (iv) Bereken de spreiding en de standaardafwijking voor het aantal sets die het match duurt: onafhankelijk van wie er wint, als A wint en als B wint.